Determination of CERES TOA Fluxes Using Machine-Learning Algorithms

Bijoy Vengasseril Thampi[1], Takmeng Wong[2]
Constantine Lukashin[2]

[1]Science Systems and Applications, Inc., Hampton, VA
[2]NASA Langley Research Center, Hampton, VA

CERES Science Team Meeting, 16-18 May 2017

Objective

- In this study, our objective is *to develop a Machine learning methodology for the determination of CERES clear scenes and subsequent clear-sky TOA flux estimation* using standalone CERES TOA radiance measurements (without any MODIS/Imager data).

Methodology

- Scene Classification - Random Forests **(RF)** method
  - Developed by Breiman and Cutler(2000)
  - Adopted for CERES –Thampi et al. (2017), submitted to JAOT
- TOA Flux estimation – Artificial Neural network **(ANN)** method
  - ANN methodology outlined in Lukashin and Loeb(2003)

Machine learning Algorithms

Random Forests (RF)	Artificial Neural networks (ANN)
○ is an ensemble learning method for classification and regression. ○ Random forests operate by constructing a multitude of decision trees and outputting the class that gets maximum number of votes from the forest. ○ Main advantages of RF method are: • *they have faster runtimes* • *can deal with unbalanced and missing data* • *has the ability to handle data without preprocessing or rescaling.*	▪ ANN is based on a large number of neural units loosely modelling the way a biological brain solves problem. ▪ They are exceptionally good at performing pattern recognition and other tasks that are very difficult to program using conventional techniques. ▪ Programs that employ neural nets are also capable of learning on their own and adapting to changing conditions.

Input data

Input Variables	IGBP surface types
Solar & viewing zenith-angles	Water bodies
Relative azimuth angle	Bright Desert
CERES Shortwave (SW) and Longwave (LW) broadband radiances	Dark Desert
LW surface emissivity	Grasslands
Broadband surface-albedo	Croplands and cities
Surface skin temperature	Evergreen Forests
Precipitable water	Deciduous Forests
Wind speed	Woody Savannas and Shrub lands
	Permanent and Fresh-snow
	Sea Ice

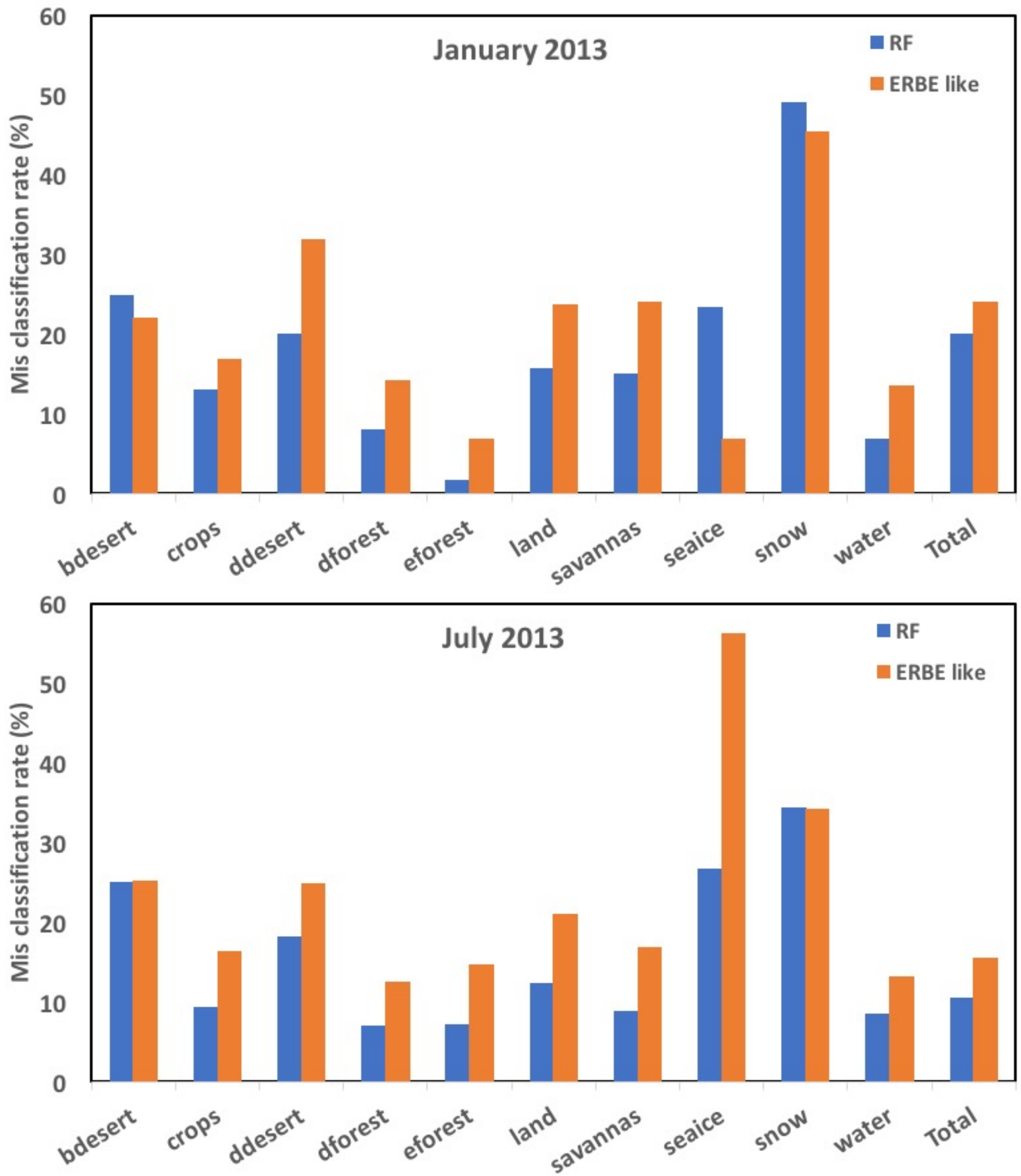
Using RF, The TOA radiances are classified in to clear and cloudy classes first.

In the second step, radiances classified as CLEAR-SKY are converted to TOA fluxes using the ANN method.

CERES Aqua SSF data

Training data : 2003-2014
Test data : 2015

Scene classification: RF vs ERBE like



Intercomparison of misclassification rate between ERBE-like and RF is carried out.

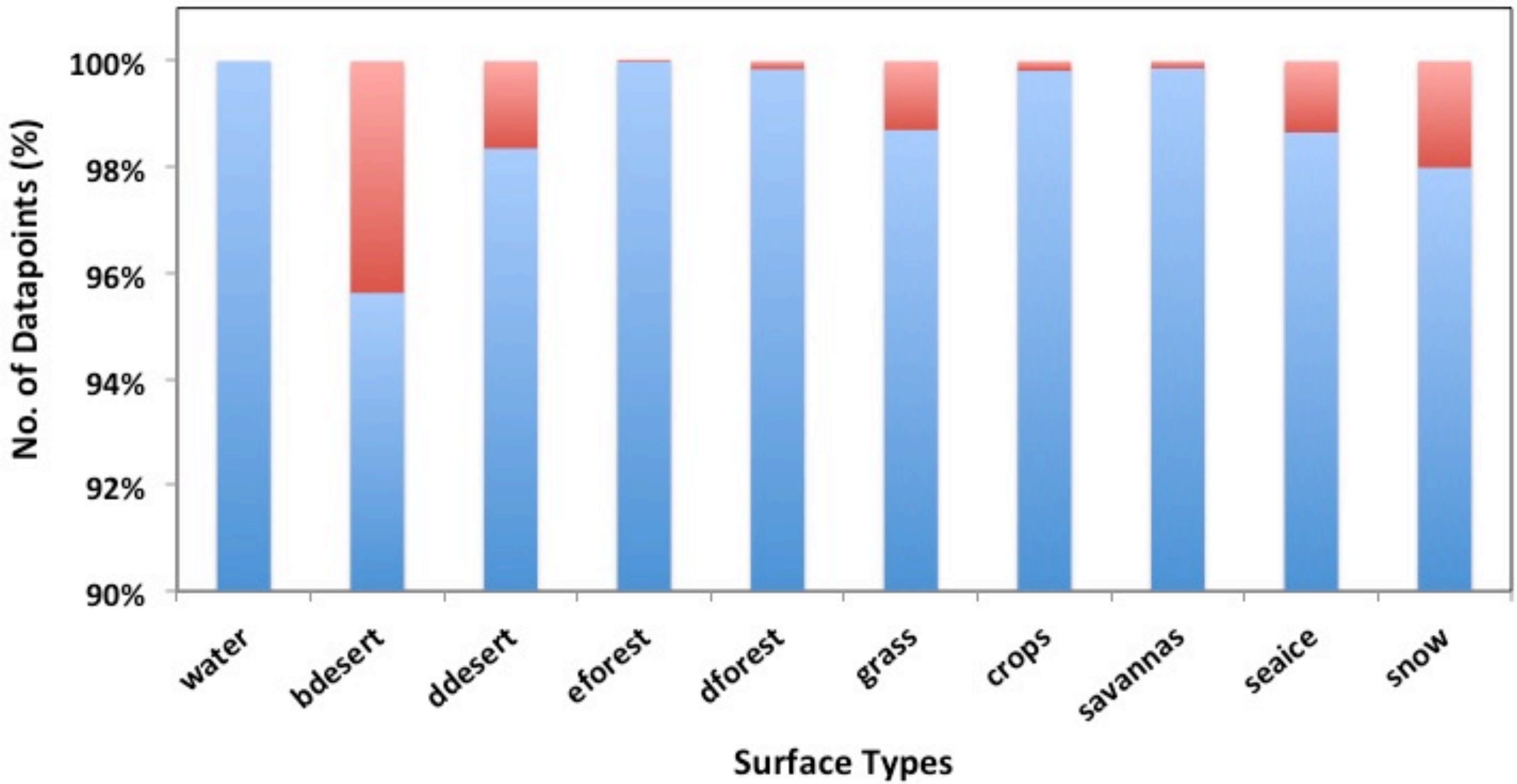
RF provides better classification for most surface types.

Snow and SeaIce surface types generally show better classification for ERBE-like data

RF Results

Month : July (Day time)

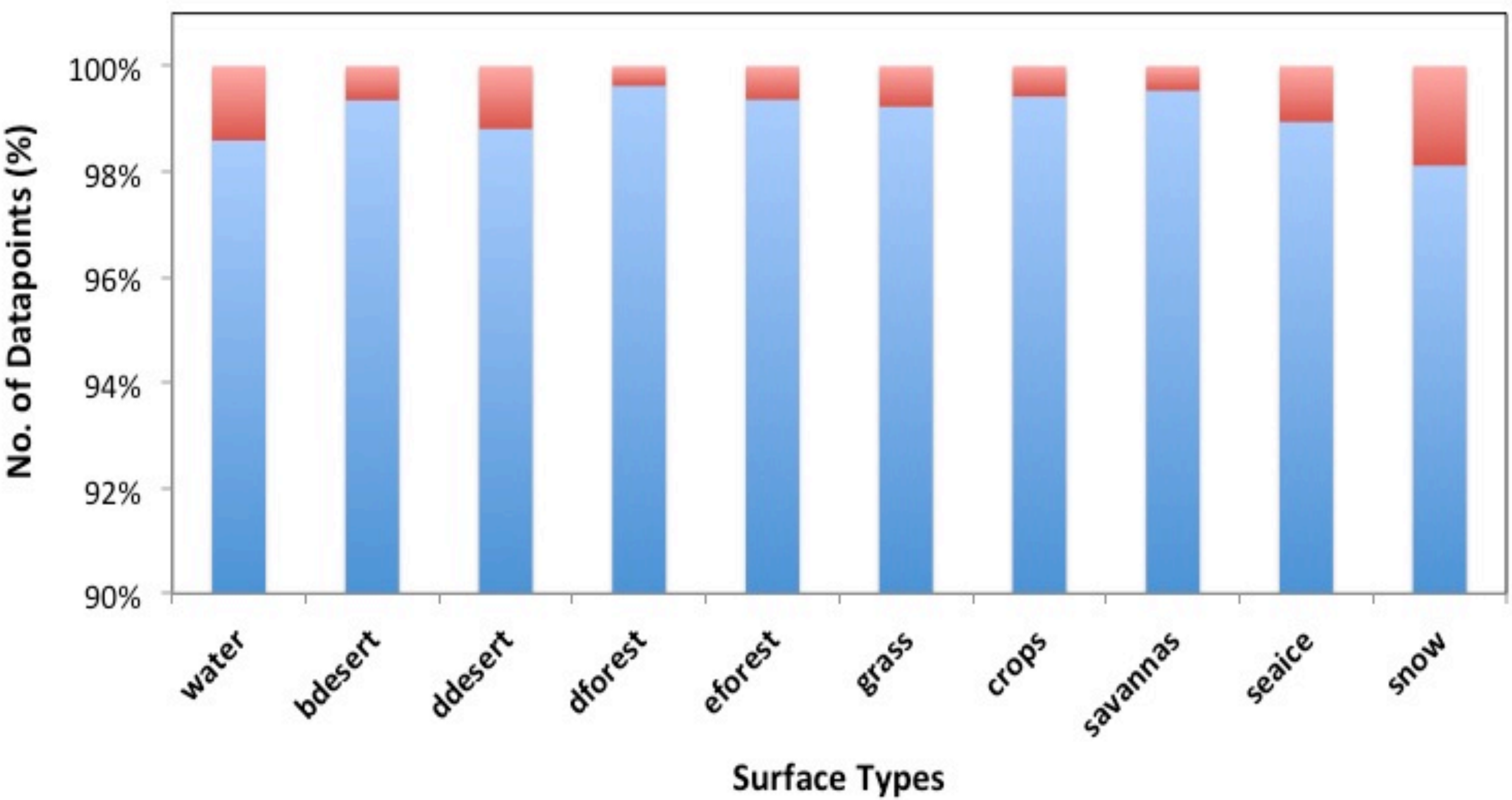
RED – misclassified data points



RF Results

Month : July (Night time)

RED – misclassified data points



ANN clear-sky Flux estimation

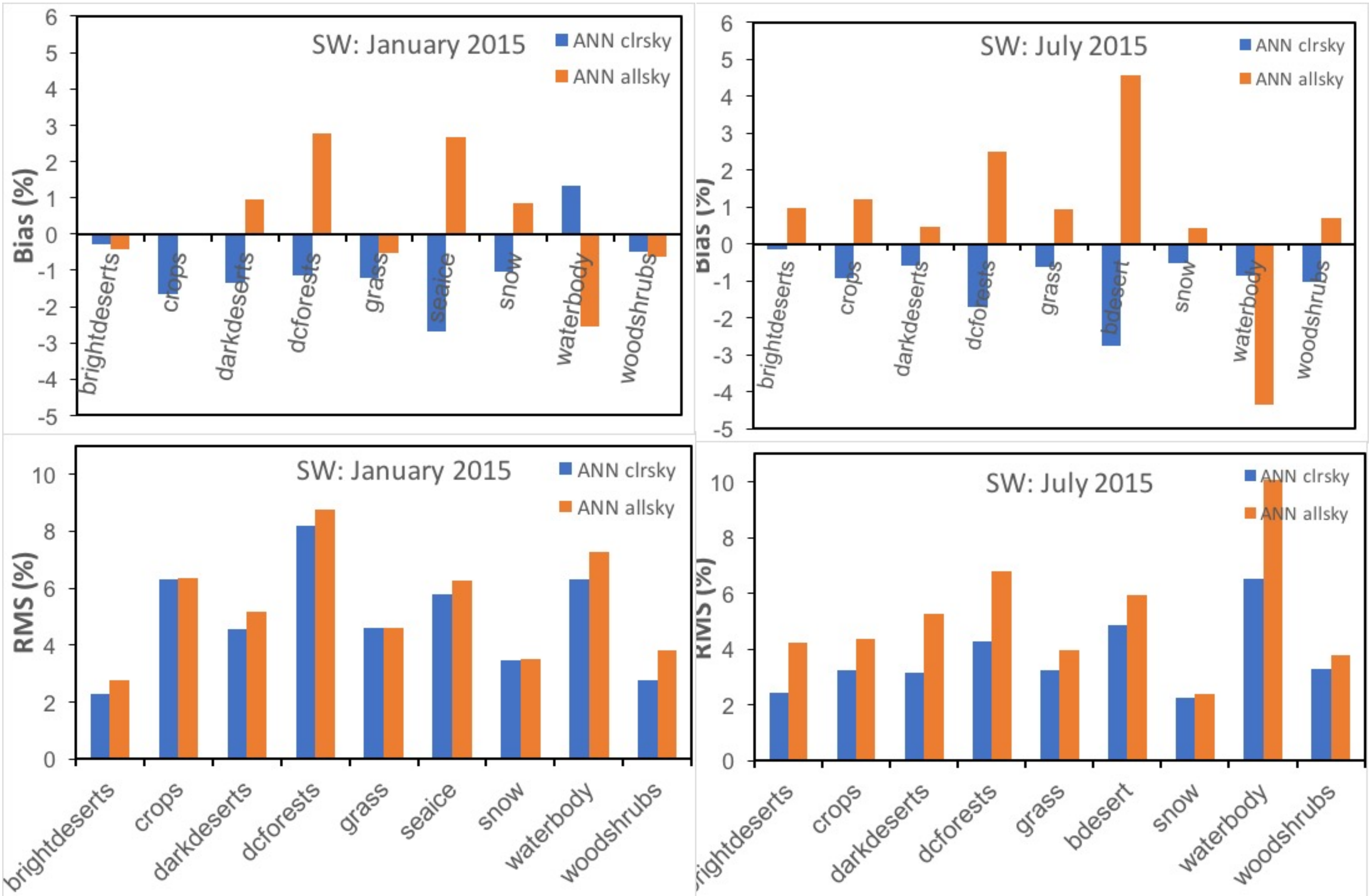
- Once the clear-scene identification is carried out by Random Forest method, CERES radiance to flux conversion is carried out by employing a feed-forward error back-propagation (**FFEB**) artificial neural network (ANN) method (**Loukachine and Loeb, 2003**).
- The technique is then validated by comparing ANN-derived TOA fluxes with CERES (original) TOA fluxes.
- In the ***modified ANN method***, only clear-sky SSF monthly data (2003-2014) is used to train the ANN and results were compared with all-sky ANN methodology.

TOA clear-sky Flux: $\text{ANN}_{\text{clear}}$ vs $\text{ANN}_{\text{allsky}}$

- ANN radiance to flux conversion of RF classified data (clear) is conducted using both modified ANN and original ANN method
- Analysis of the ANN derived Flux show that ANN clear sky method produce better results for majority of the cases (>60%) compared to the ANN all sky method.

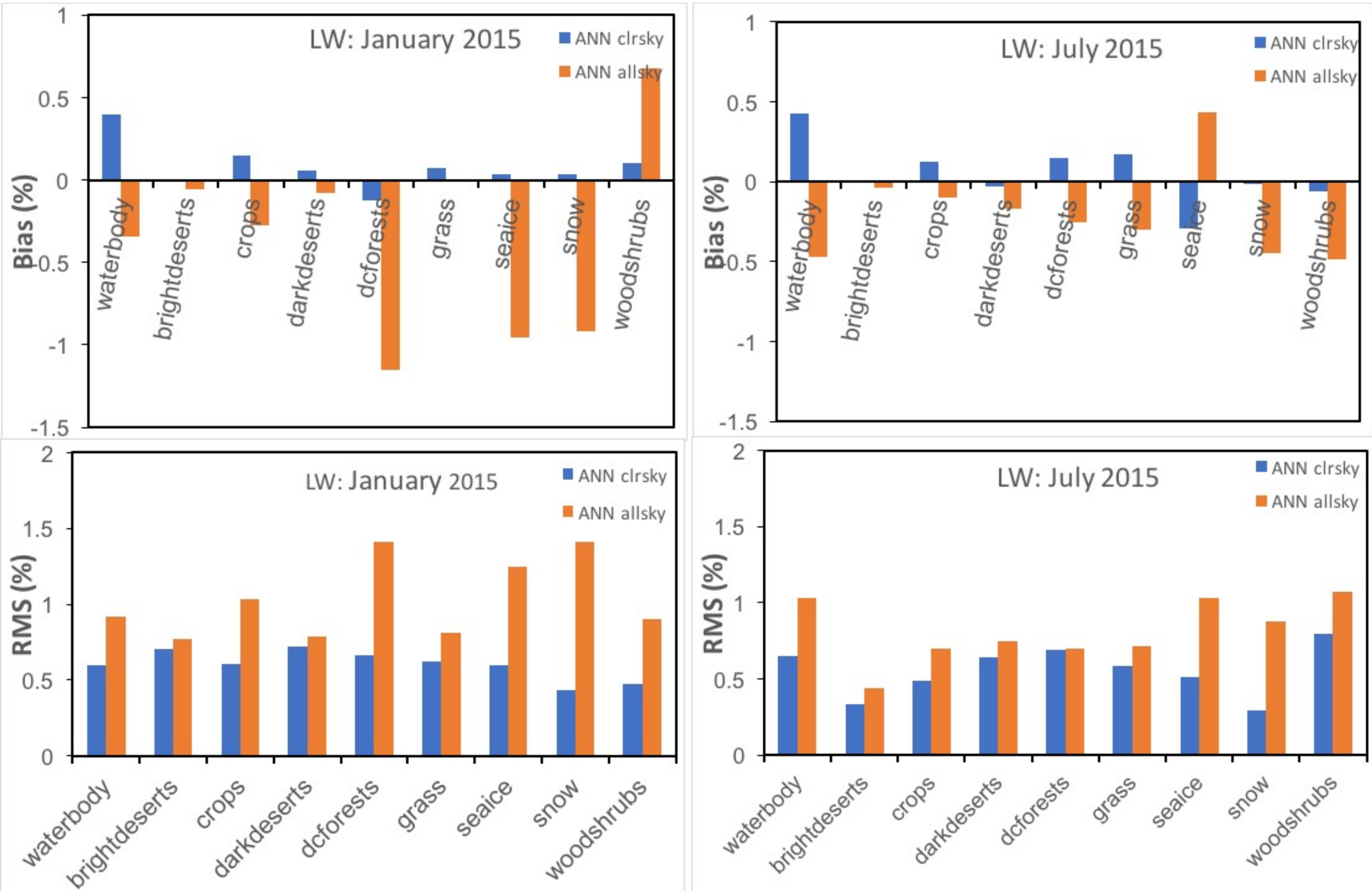
SURFACE TYPE	SW		LW	
	JAN (%)	JUL(%)	JAN(%)	JUL(%)
bdesert	64.5	67.3	84.1	63.7
crops	59.2	63.6	85.4	88.8
ddesert	57.3	64.7	82.8	77.1
dforest	65.0	68.6	63.7	59.8
grass	65.5	73.9	80.4	49.5
savannas	62.2	74.3	59.2	61.8
seaice	62.4	68.6	76.0	68.9
snow	63.5	77.4	60.9	71.2
water	58.1	67.9	67.4	67.0

Bias & RMS : SW Clear-sky Flux



$$\text{Bias} = \text{Flux}_{ANN} - \text{Flux}_{SSF}$$

Bias & RMS : LW Clear-sky Flux



Summary

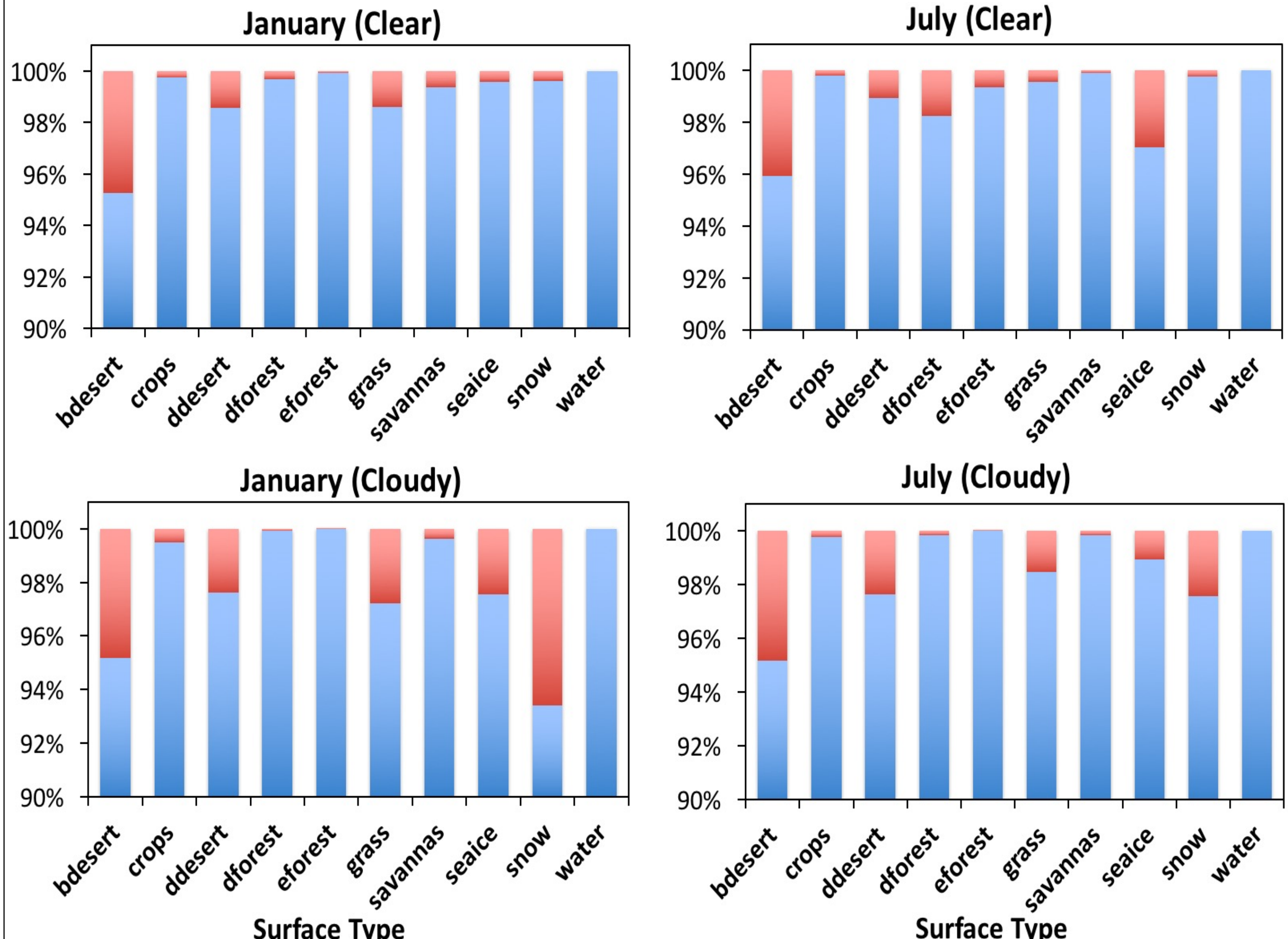
- A new methodology for Imager independent CERES TOA clear-sky flux retrieval is developed incorporating ***Random Forests*** scene classification and ***Artificial Neural Network*** flux estimation methods.

- RF misclassification rate for (Clear and cloudy, Day time) shows lower values (< 2%) for Water bodies, Crops, Evergreen forest, etc.

- Modified ANN clear-sky method produce more accurate TOA flux values most of the time (>60% of data) compared to all-sky ANN method with relatively lower Bias.

Thank you…

RF scene classification Results

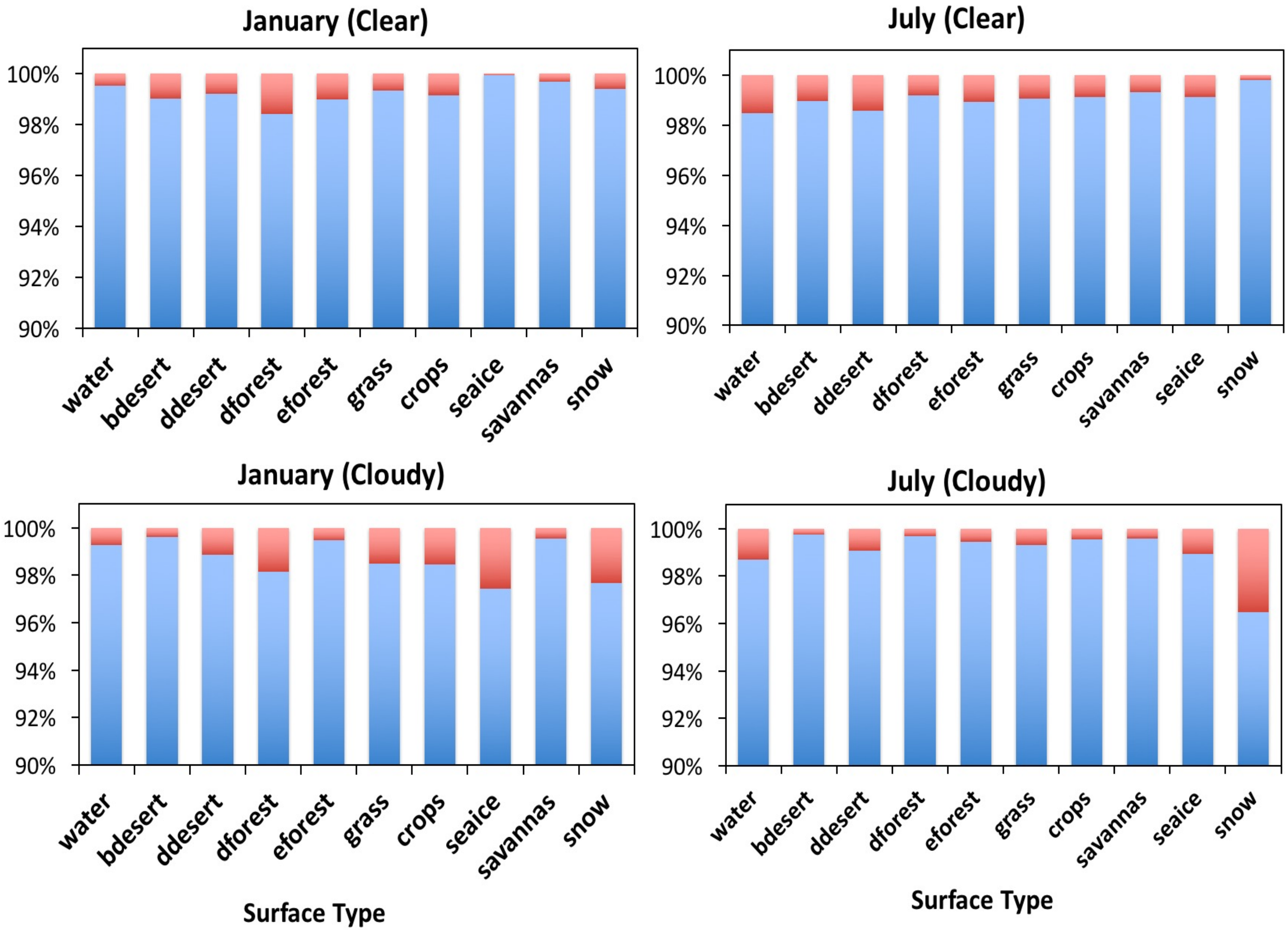


Year : 2015
(Day time)

RED – misclassified data points

- Scene Classification rate in general is > 98% for most of the surface types
- A misclassification rate of 3-6% is observed for surface types like bright deserts, snow and seaice.

RF scene classification Results

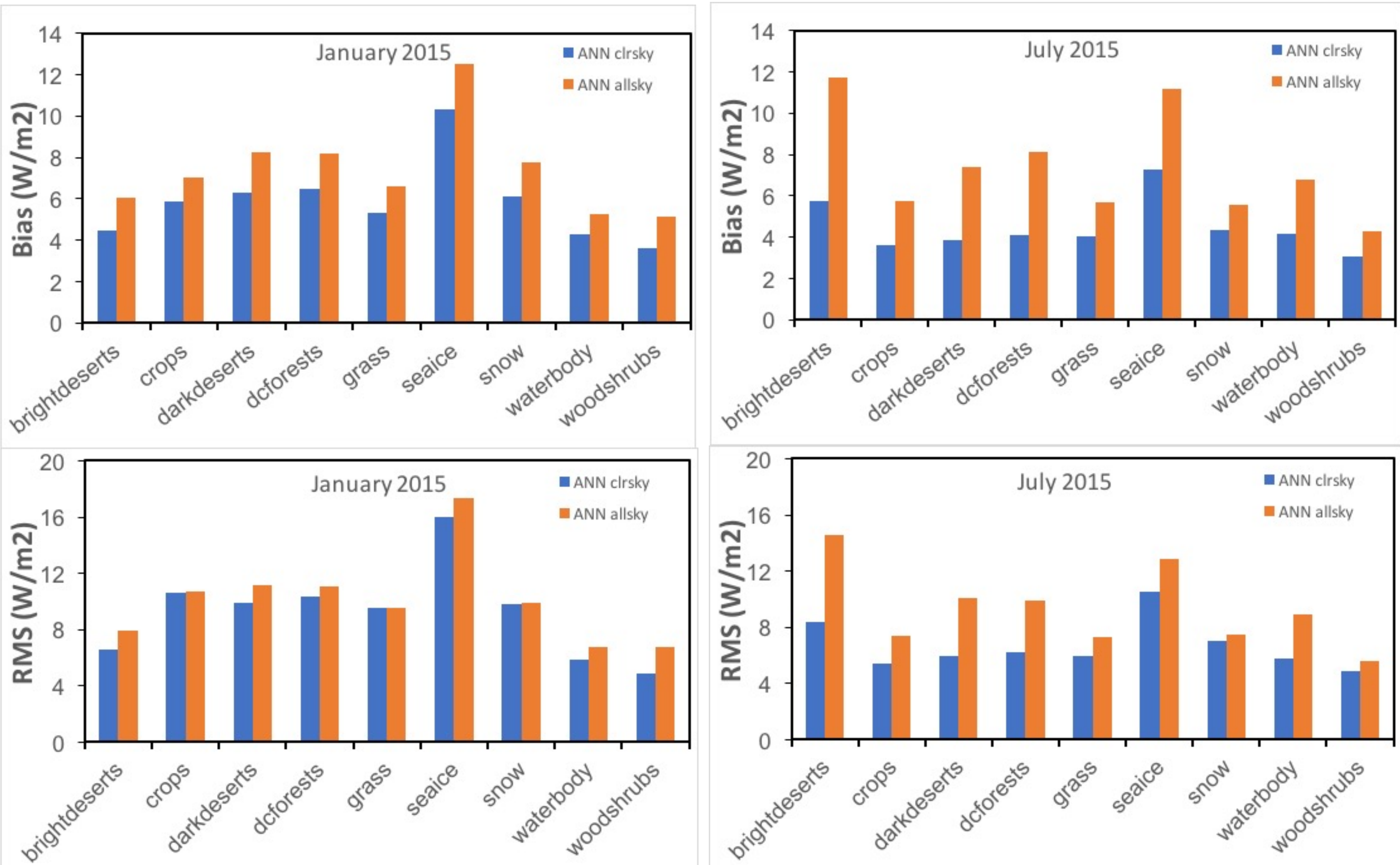


Year : 2015 (Night time)

RED – misclassified data points

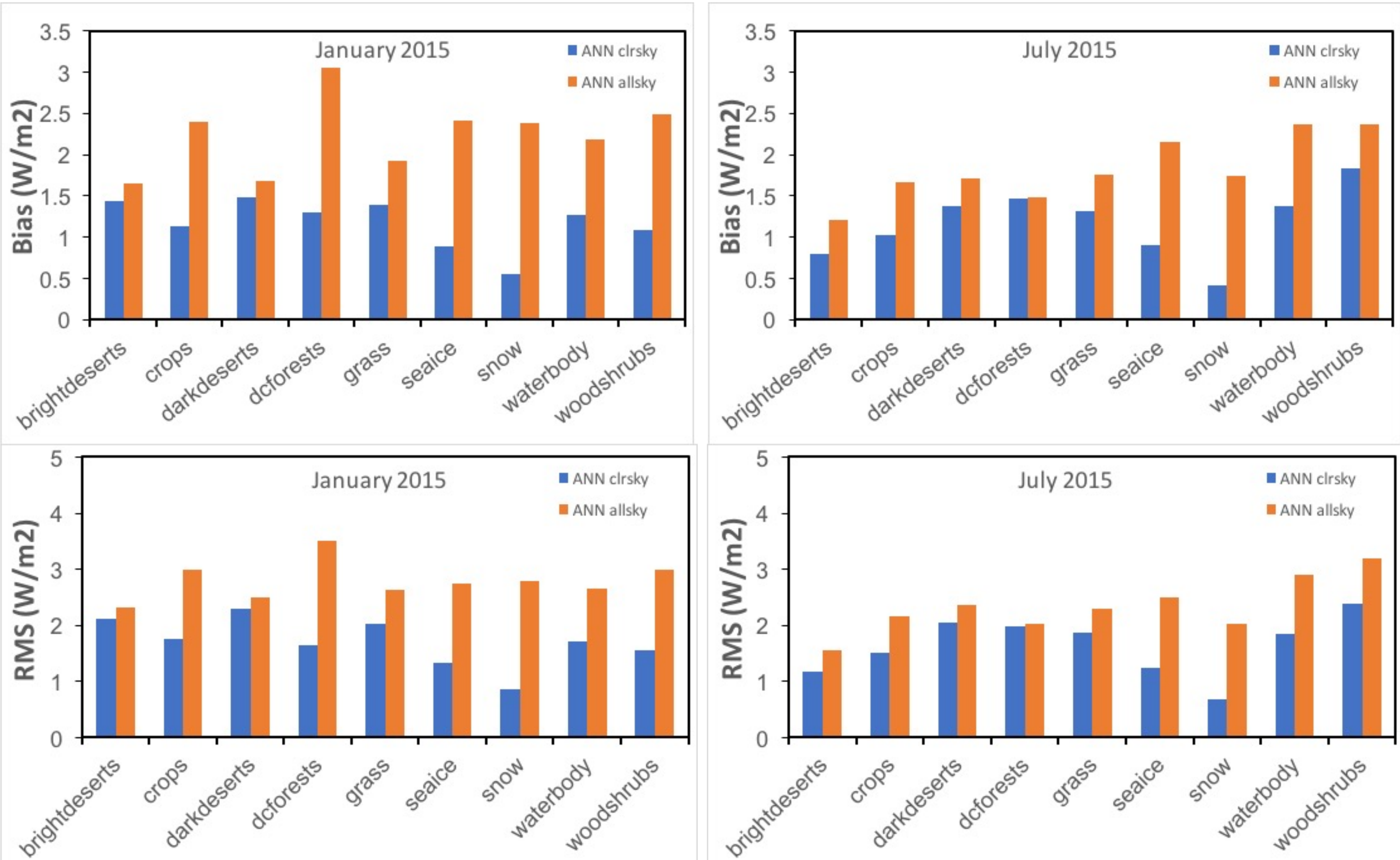
- Both SW radiance and albedo are not included in the night time analysis
- Scene classification rate in general > 98% for most of the surface types
- Misclassification rate is relatively high (>3%) for surface types like snow and seaice.

Absolute Bias & RMS : SW clear-sky Flux

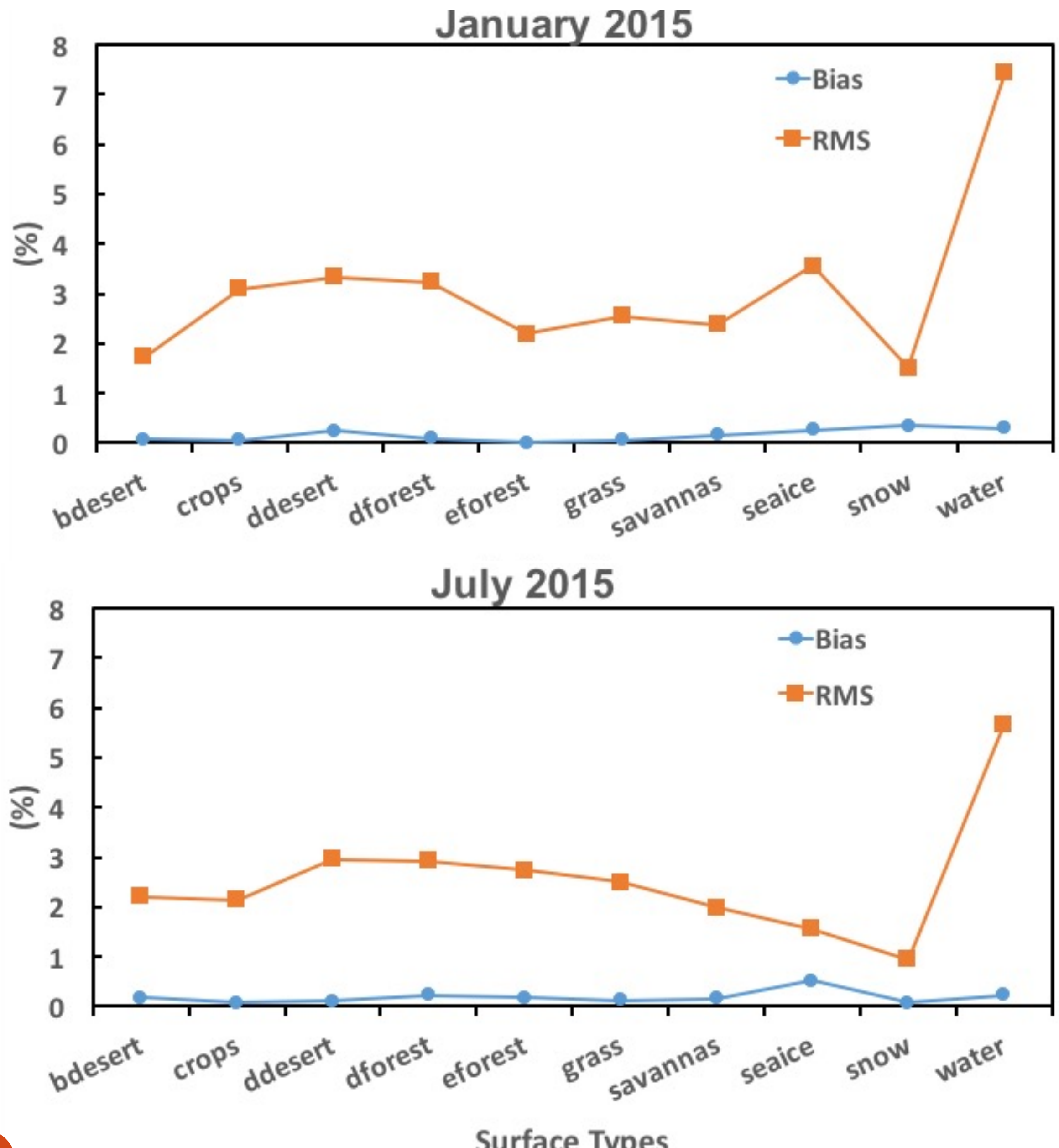


- Mean Bias and RMS is relatively lower for the ANN clear sky method compared to the all sky method estimated for the Clear-sky SW TOA Fluxes.

Absolute Bias & RMS : LW clear-sky Flux



Global mean BIAS and RMS : Shortwave flux



- Global mean Bias in SW flux on the other hand is lower (< 1%) for all the surface types

- Mean RMS in SW Flux (%) for most surface types are below 4% while it is relatively higher (6-8%) for water surface.

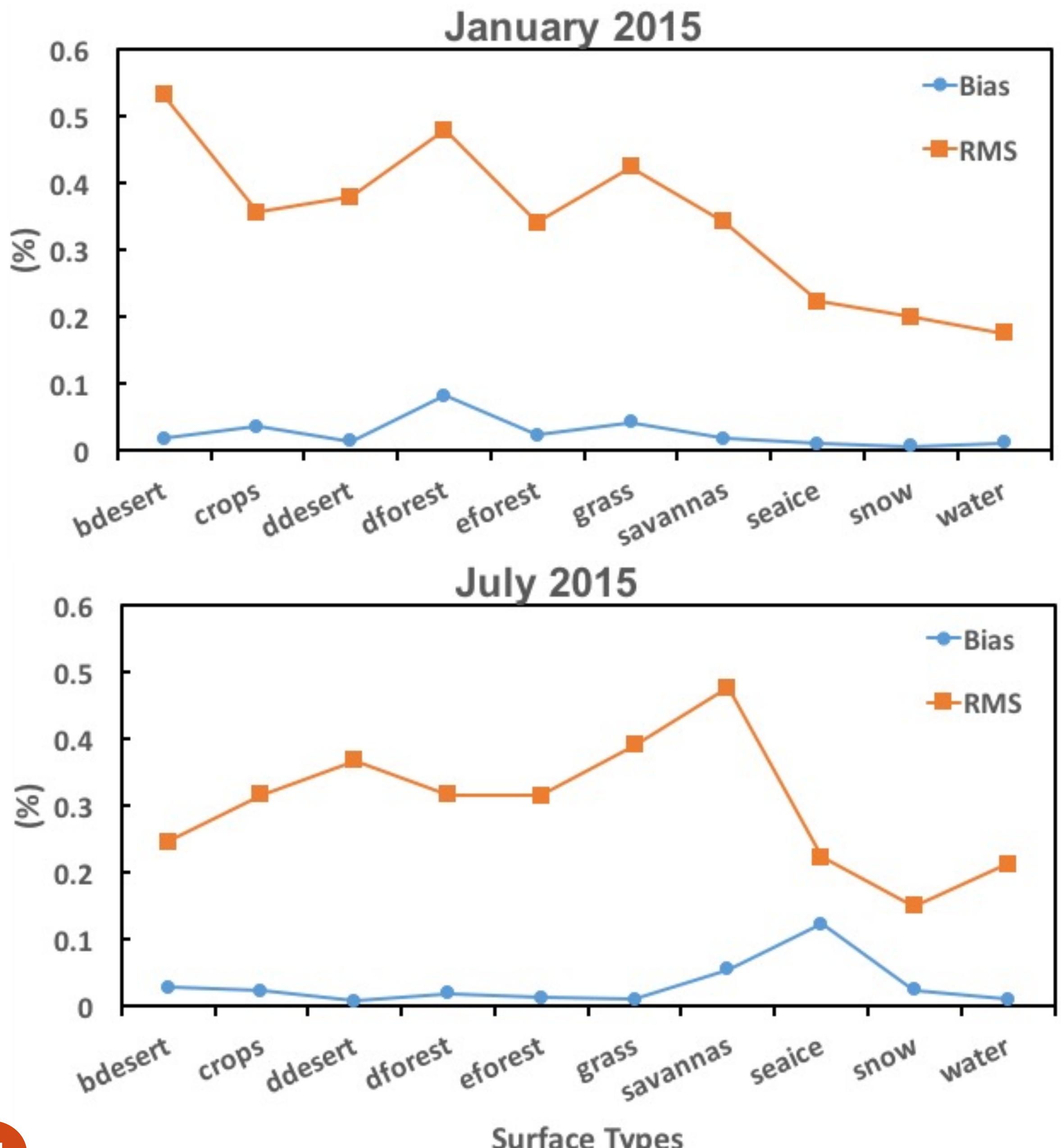
Random Forests- Training data

Input variables are selected for the scene classification are:

CERES	Ancillary data
Solar zenith & viewing zenith angles	LW surface emissivity
Relative azimuth angle	Broadband surface albedo
CERES LW & SW broadband radiances	Surface skin temperature
IGBP Surface type	Precipitable water
	Wind speed

IGBP Surface Types	
Water bodies	Evergreen Forests
Bright Desert	Deciduous Forests
Dark Desert	Woody Savannas and Shrub lands
Grasslands	Permanent and Fresh snow
Croplands and cities	Sea Ice

Global mean BIAS and RMS : Longwave flux



- Similarly, the global mean Bias for the LW flux is also lower (< 0.15%) for all the surface types compared to the SW flux values.
- Compared to the SW flux, Mean RMS for the LW Fluxes are usually below 1%